

Toward Domain-Specific Russian-Tatar Thesaurus Construction

Galieva A., Kirillovich A., Khakimov B., Suleymanov D., Nevzorova O., Loukachevitch N.
Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2017 Association for Computing Machinery. The paper discusses the main principles and practical aspects of implementing a new bilingual lexical resource - the Russian-Tatar thesaurus on socio-political and IT issues. This thesaurus is developed on the basis of the Russian RuThes thesaurus format which is built as a hierarchy of concepts viewed as units of thought, with each concept linked to a set of language expressions that refer to it in texts (text entries). The paper discusses general methodology of translating concept names and their text entries, as well as ways of reflecting the specificity of the Tatar lexical-semantic system.

<http://dx.doi.org/10.1145/3143699.3143716>

Keywords

IT terminology concept, Socio-political terminology, Text entry, The Tatar language, Thesaurus, Vocabulary

References

- [1] Amirov, F.K. 1996. Russian-Tatar Legal Dictionary [Russko-tatarskij yuridicheskij slovar']. Kazan, Russia.
- [2] Galieva, A.M., Kirillovich, A.V., Loukachevitch, N.V., Nevzorova, O.A., and Suleymanov D.Sh. 2017. Creating Russian-Tatar Socio-Political Thesaurus: General Principles and Implementation [Sozdanie rusko-tatarskogo tezaurusa po obshchestvenno-politicheskoy tematike: obshchie principy i aspekty realizacii'. In Scientific and Technical Information. Series 2: Information Processes and Systems [Nauchno-tekhnicheskaya informaciya. Seriya 2: Informacionnye processy i sistemy], 2, 20-28.
- [3] GOST 7.24-2007 Interstate Standard. Multilingual Thesaurus of information retrieval [Mezhgosudarstvennyj standart. Tezaurus informacionno-poiskovyj, mnogoyazychnyj]. Moskva, Standartinform.
- [4] INIS Thesaurus. English version. IAEA-INIS Reference Series. 2016.
- [5] ISO 25964-1:2011. Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval. URL: http://www.iso.org/iso/catalogue_detail.htm?csnumber=53657.
- [6] ISO 5964-1985. Documentation – Guidelines for the establishment and development of multilingual thesauri. URL: http://www.iso.org/iso/catalogue_detail.htm?csnumber=12159.
- [7] Khakimov, B.E. 2008. Some Aspects of Transplantation of Computer Science Texts (basing on the example of Tatar Localization of Windows OS) [O nekotoryh voprosah perevoda komp'yuterno-tekhnicheskikh tekstov (na primere tatarskoj lokalizacii OS Windows). In Philological Sciences. Theoretical and Practical Issues [Filologicheskie nauki. Voprosy teorii i praktiki]. Vol. 2, 134-136.
- [8] Linguistic Linked Open Data Cloud (LLOD). 2015. URL: <http://linguistic-lod.org/lod-cloud>.
- [9] Loukachevitch, N. and Dobrov B. 2015. The Sociopolitical Thesaurus as a resource for automatic document processing in Russian. In Terminology. Vol. 21, 2, 237-262.
- [10] Miller, G.A. 1995. WordNet: A lexical database for English. In Communications of the ACM. Vol. 38, 11, 39-41.

- [11] MultiWordNet. 2004. URL: <http://multiwordnet.fbk.eu/english/home.php>.
- [12] Nizamov, I.M. 1995. Brief Russian-Tatar Socio-Political Dictionary [Kratkij russko-tatarskij obshchestvenno-politicheskij slovar]. Kazan, Russia.
- [13] Nizamov, I.M., et al. 1997. Russian-Tatar Socio-Political Dictionary [Russko-tatarskij obshchestvenno-politicheskij slovar']. Kazan, Russia.
- [14] Suleymanov, D., Galieva, A., Galimyanov, A. 2016. Formation and use of Tatar Computer Science Terms [K probleme obrazovaniya i ispol'zovaniya tatarskoj terminologicheskoy sistemy v informatike i informacionnyh tekhnologiyah]. In Proceedings of 4 International Forum on Cognitive Modeling [Kognitivnoe modelirovanie: Trudy Chetvertogo Mezhdunarodnogo foruma po kognitivnomu modelirovaniyu], 441-449.
- [15] Suleymanov, D.Sh. 2015, ed. English-Russian-Tatar-Chuvash Dictionary of Terms on Computer Science and Information Technologies [Anglo-russko-tatarsko-chuvashskij slovar terminov po informatike i informacionnym tekhnologiyam (s tolkovaniyami na tatarskom yazyke)]. Kazan, Russia.
- [16] United Nations. UNBIS Thesaurus. English version. 1986. New York, Dag Hammarskjold Library of the United Nations.
- [17] Fellbaum C. 1998, ed. WordNet: An Electronic Lexical Database. 1998. ed., Cambridge, MA: MIT Press.